PARTICLE TRACKING BASED ON FULLY UNSUPERVISED DISENTANGLEMENT OF THE GEOMETRICAL FACTORS OF VARIATION

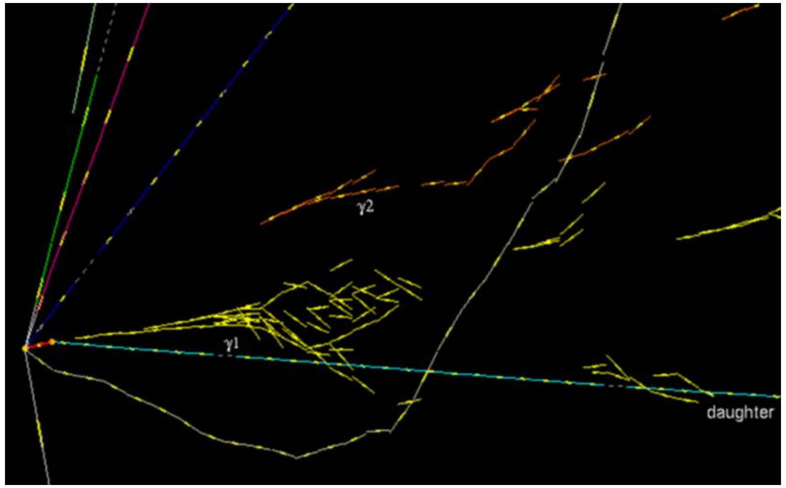
Mykhailo Vladymyrov

Laboratory for High Energy Physics
University of Bern, Switzerland

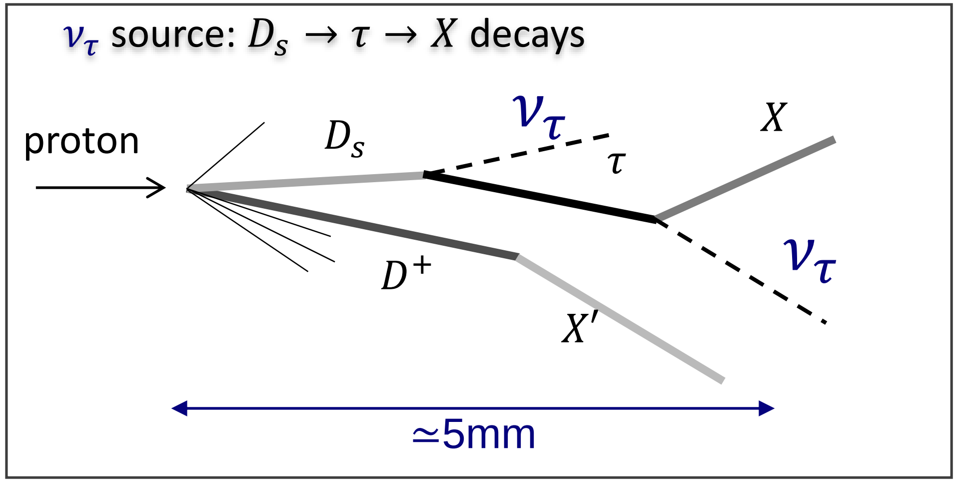
Currently at the Theodor Kocher Institute & Science IT Support, University of Bern

Particle tracking in neutrino physics

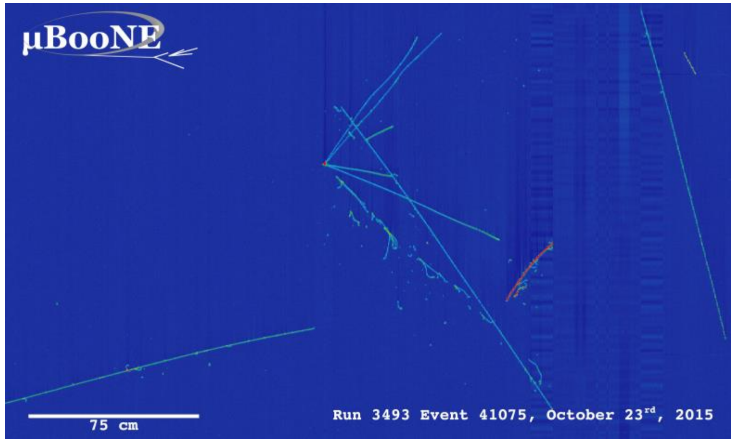
ν_τ Appearance in OPERA

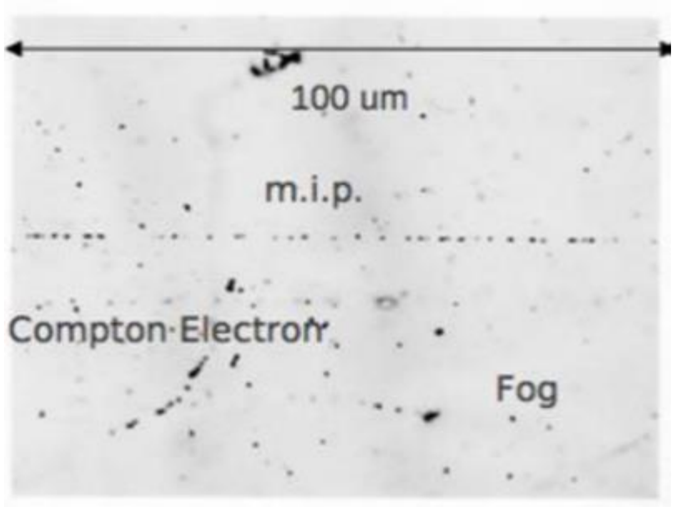


ν_τ production study in DsTau (NA65)

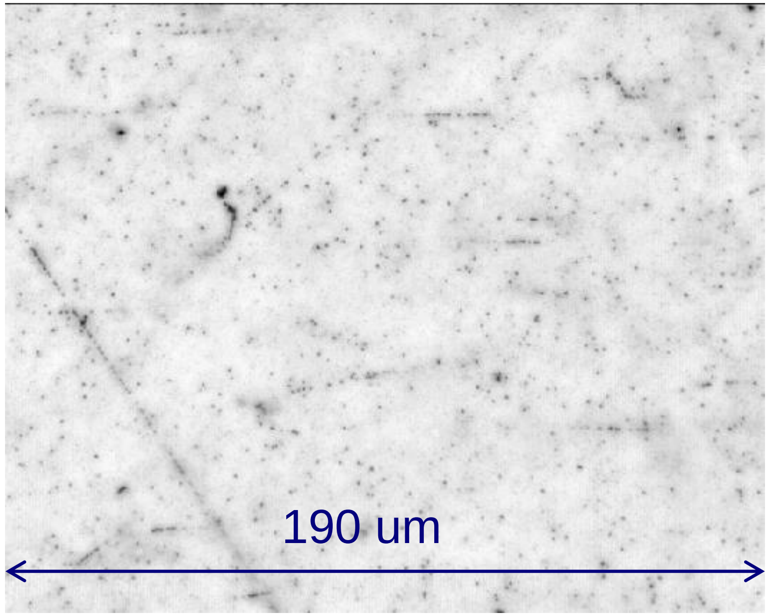


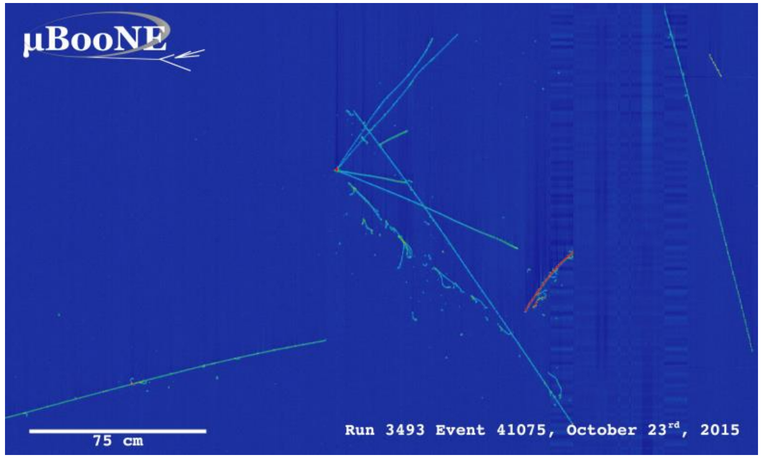
First ν interaction in μBooNE





Tracking: finding lines





Algorithms:

- Leverage small part of available information
- Tedious adaptation to new conditions

Supervised Machine Learning

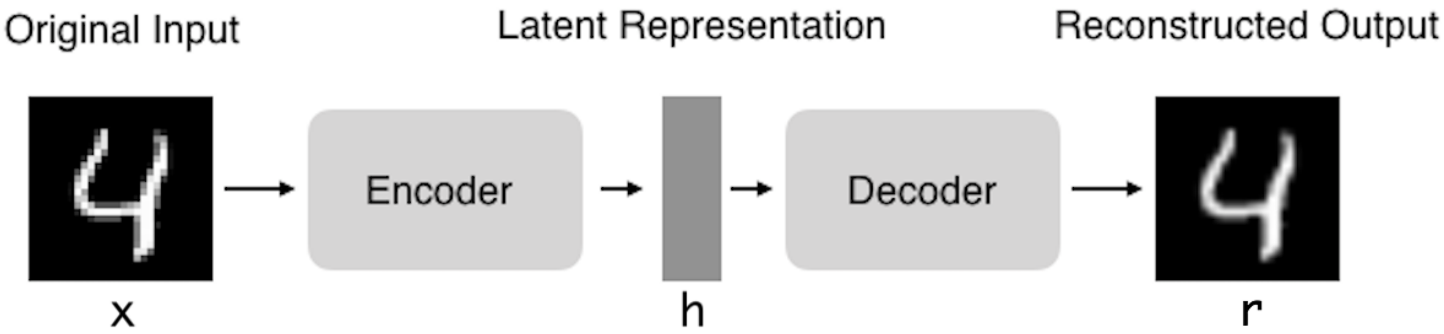
- Leverage all low-level information
- Requires labeled data / training on simulations

Unsupervised Machine Learning

- Leverage all low-level information
- No / small amount of labeled data

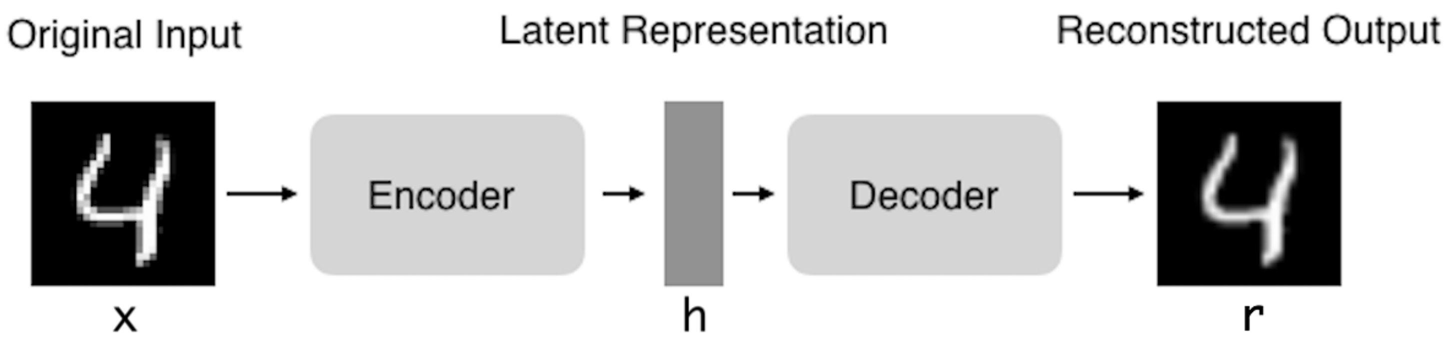
Autoencoder

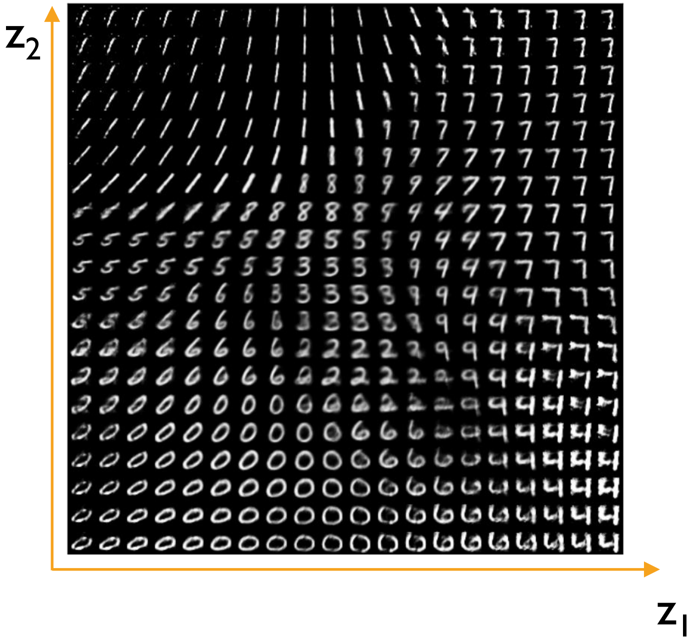
$$\phi : \mathcal{X} \rightarrow \mathcal{F}$$
$$\psi : \mathcal{F} \rightarrow \mathcal{X}$$
$$\phi, \psi = \underset{\phi,\psi}{\arg\min} \, \|X - (\psi \circ \phi)X\|^2$$



Autoencoder

$$\phi : \mathcal{X} \rightarrow \mathcal{F}$$
$$\psi : \mathcal{F} \rightarrow \mathcal{X}$$
$$\phi, \psi = \underset{\phi,\psi}{\arg\min} \, \|X - (\psi \circ \phi)X\|^2$$





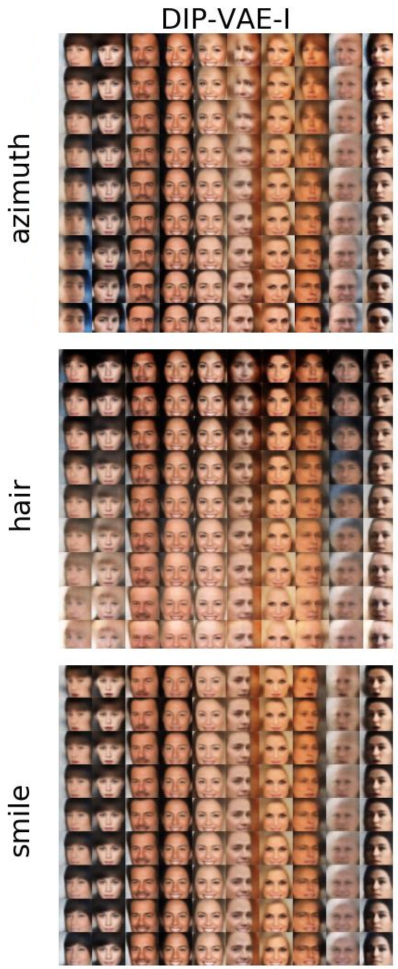
How to map latent representation to meaningful parameters?

Disentanglement of latent variables

Independent parameters can be often disentangled in the latent representations.

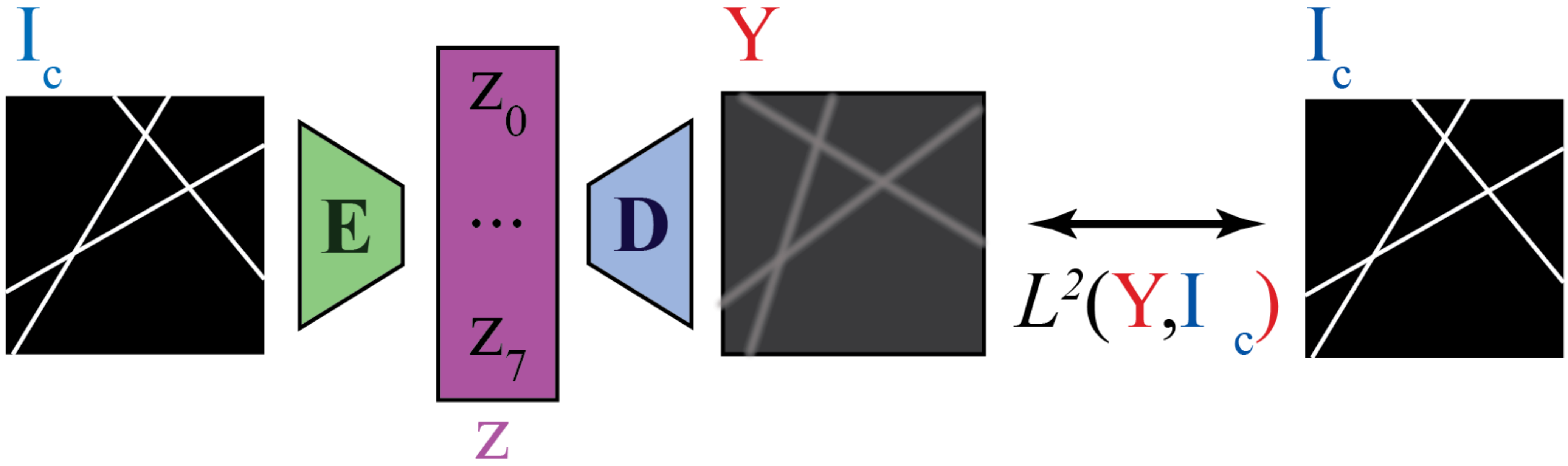
For example by:

- Distribution
- Mixing parameters
- etc



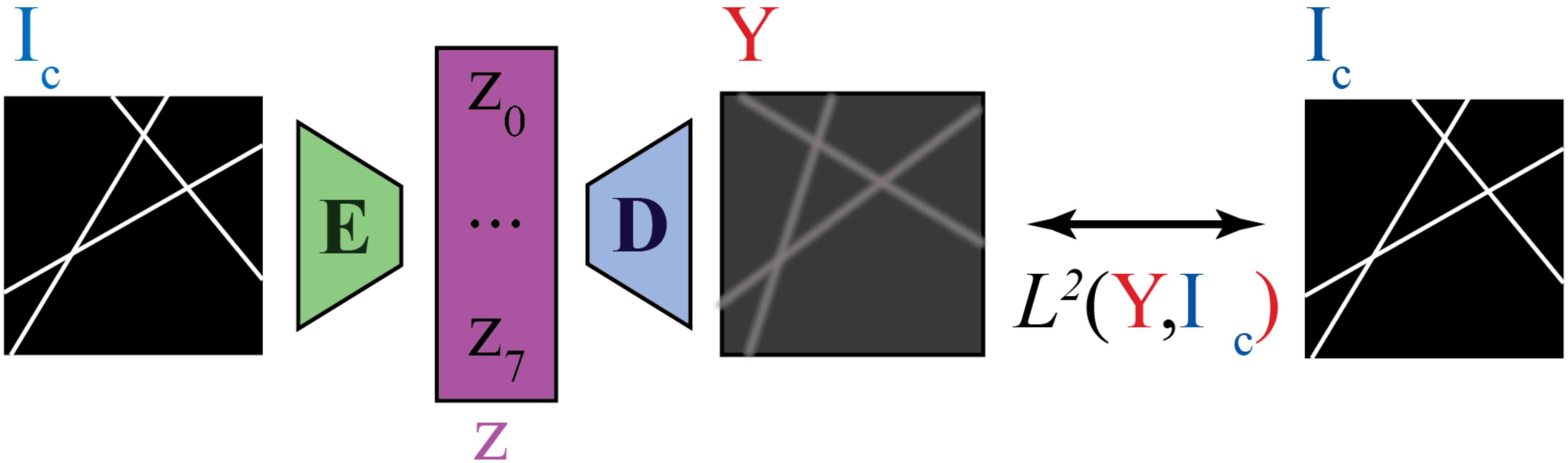
A. Kumar et al, 2017. Variational Inference of Disentangled Latent Concepts from Unlabeled Observations.

Proposed model



z_i – parameters that *shall* correspond to the i^th track

Proposed model



z_i – parameters that *shall* correspond to the i^{th} track

We need to 1. disentangle different tracks; 2.disentangle track parameters

Equivariance constraint

$I' = T_{im}(I|\xi),\ z'_t = T_{tr}(z_t|\xi)$ – geometrical transformations in the image and the latent representation space, parametrized by the same parameter set ξ.

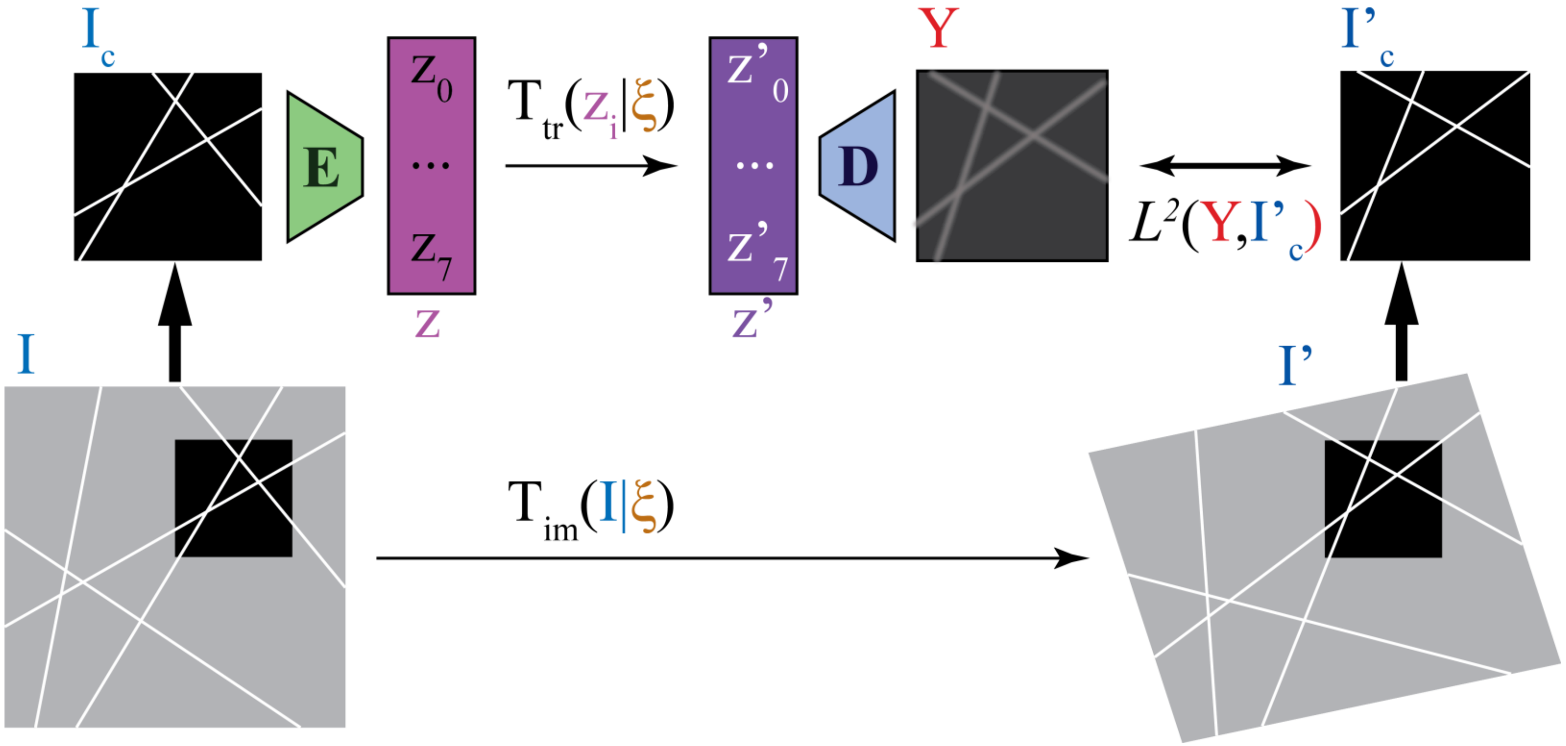
Demand equivariance of encoder and decoder:

$$T_{tr}(E(I)|\xi) = E(T_{im}(I|\xi));\ \forall\xi$$
$$D\big(T_{tr}(z|\xi)\big) = T_{im}(D(z)|\xi);\ \forall\xi$$

->

$$D\big(T_{tr}(E(I)|\xi)\big) = T_{im}(I|\xi);\ \forall\xi,$$

Equivariance constraint



$I' = T_{im}(I|\xi)$, $z'_t = T_{tr}(z_t|\xi)$ – geometrical transformations in the image and the latent representation space, parametrized by the same parameter set ξ.

Demand equivariance of encoder and decoder:

$$T_{tr}(E(I)|\xi) = E(T_{im}(I|\xi));\ \forall\xi$$
$$D\big(T_{tr}(z|\xi)\big) = T_{im}(D(z)|\xi);\ \forall\xi$$

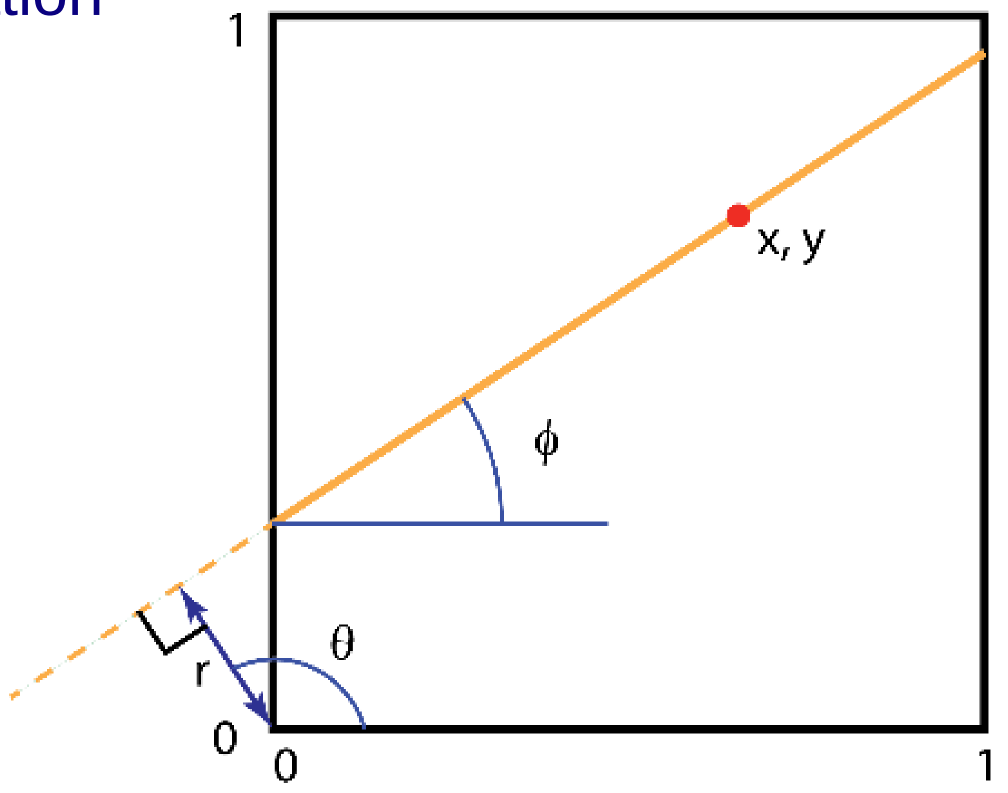
->

$$D\big(T_{tr}(E(I)|\xi)\big) = T_{im}(I|\xi);\ \forall\xi,$$

$$E,D = \underset{E,D}{\operatorname{argmin}}\ \underset{I,\xi}{\mathbb{E}}\, L^2(Y, I'_c) = \underset{E,D}{\operatorname{argmin}}\ \underset{I,\xi}{\mathbb{E}}\, L^2\left(D\left(T_{tr}\big(E(C(I))\big|\xi\big)\right), C\big(T_{im}(I|\xi)\big)\right).$$

Latent representation

- Lines have translation invariance along itself
- Lines remain lines under Affine transformation



2D, 32x32 pixel image crops

Track line is parametrized by

— point on the track (x, y)

— slope angle ϕ as $c = \kappa\cos(\phi)\,, s = \kappa\sin(\phi)$

8 tracks parameter containers.

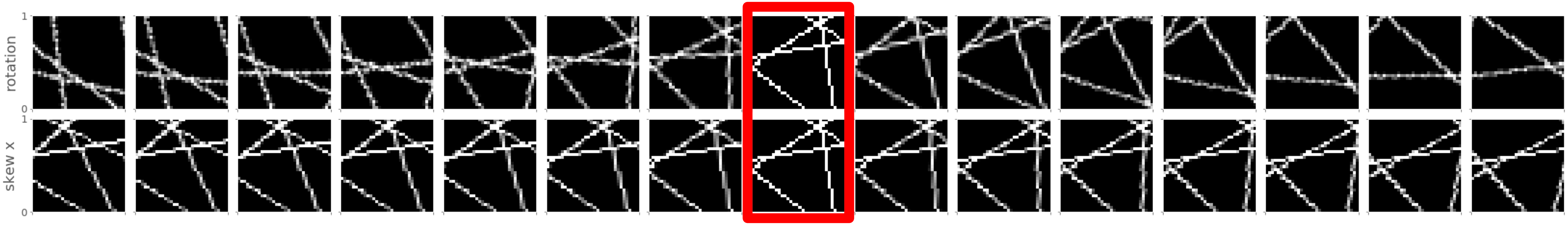
Representation transformations

Name	Description	n_p	Parametrization
RT+TI+A	Translation Invariance + Rotation transformation only with track activation parameter a	5	$(x_i, y_i, c_i, s_i, a_i)$
AT+TI	Affine Transformations + Translation Invariance	4	(x_i, y_i, c_i, s_i)

Representation transformations

Name	Description	n_p	Parametrization
RT+TI+A	Translation Invariance + Rotation transformation only with track activation parameter a	5	$(x_i, y_i, c_i, s_i, a_i)$
AT+TI	Affine Transformations + Translation Invariance	4	(x_i, y_i, c_i, s_i)

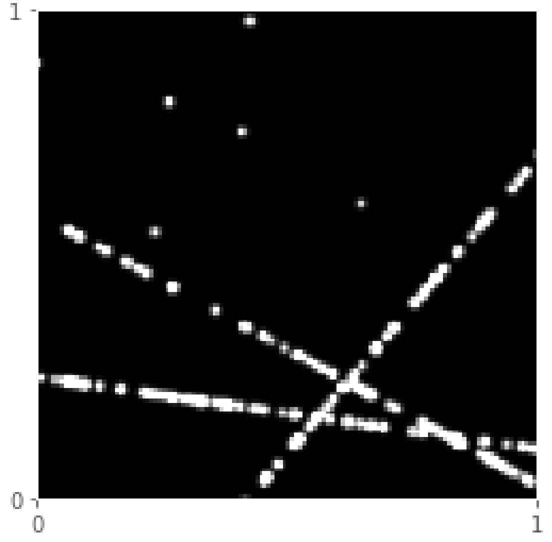
Affine Transformations: rotation, translation, skew, scale



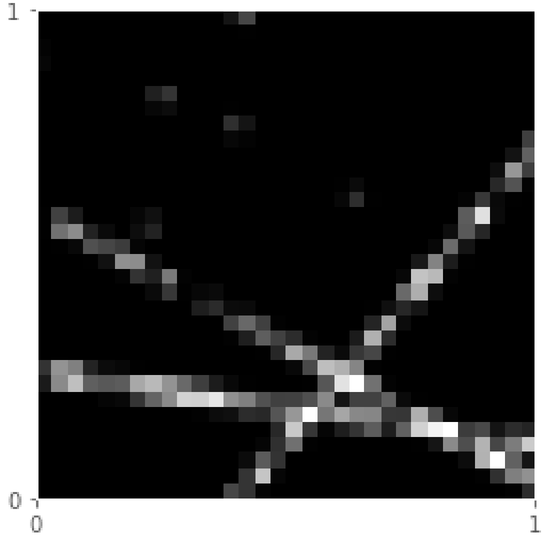
Translation invariance: translation along the line:

$$T_{t.i.}(z_t) \;=\; (x + r\,c, y + r\,s, c, s);\; r = rand(-0.5, 0.5)$$

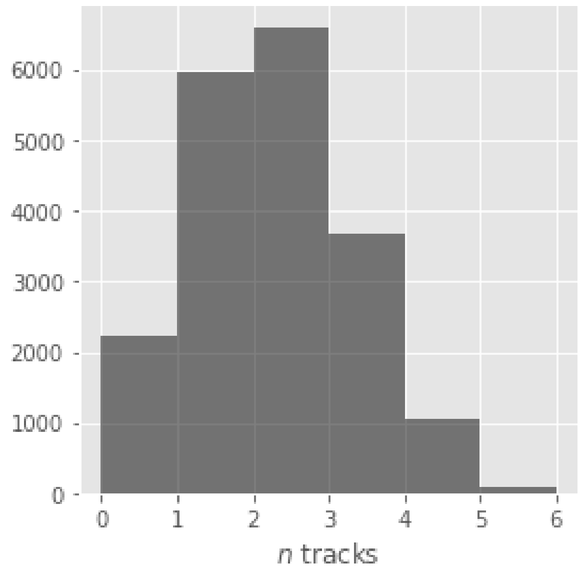
Synthetic training data

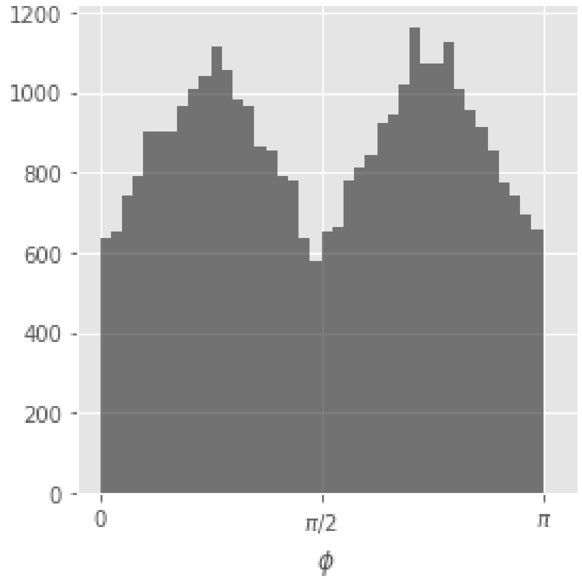


Generated dataset sample.



Downscaled sample used for model training





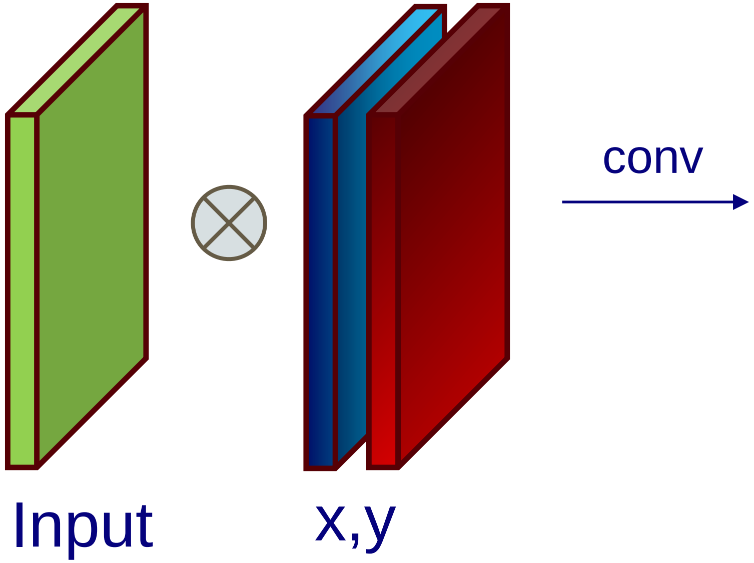
Distribution of track number per image and track angles.

Model implementation

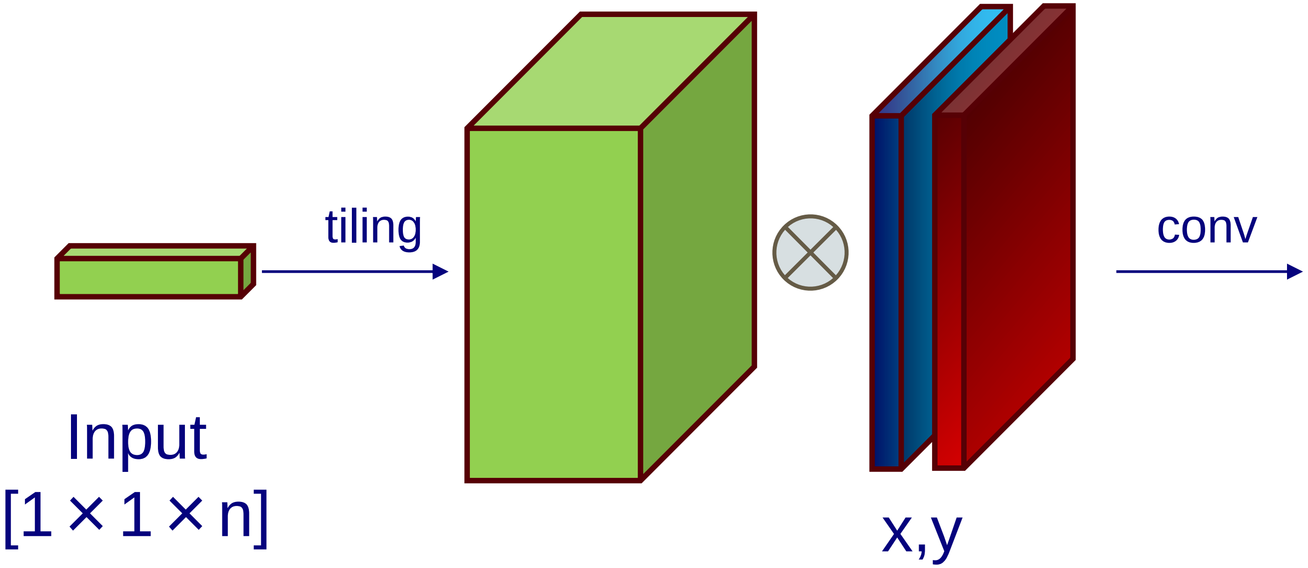
Deep convolutional autoencoder:

- Encoder: 1 CoordConv, 7 Conv, 4 Fully Connected
- Decoder: 1 Fully Connected, 1 Transposed CoordConv, 5 Conv

CoordConv: concatenation with x,y coordinate map, convolution

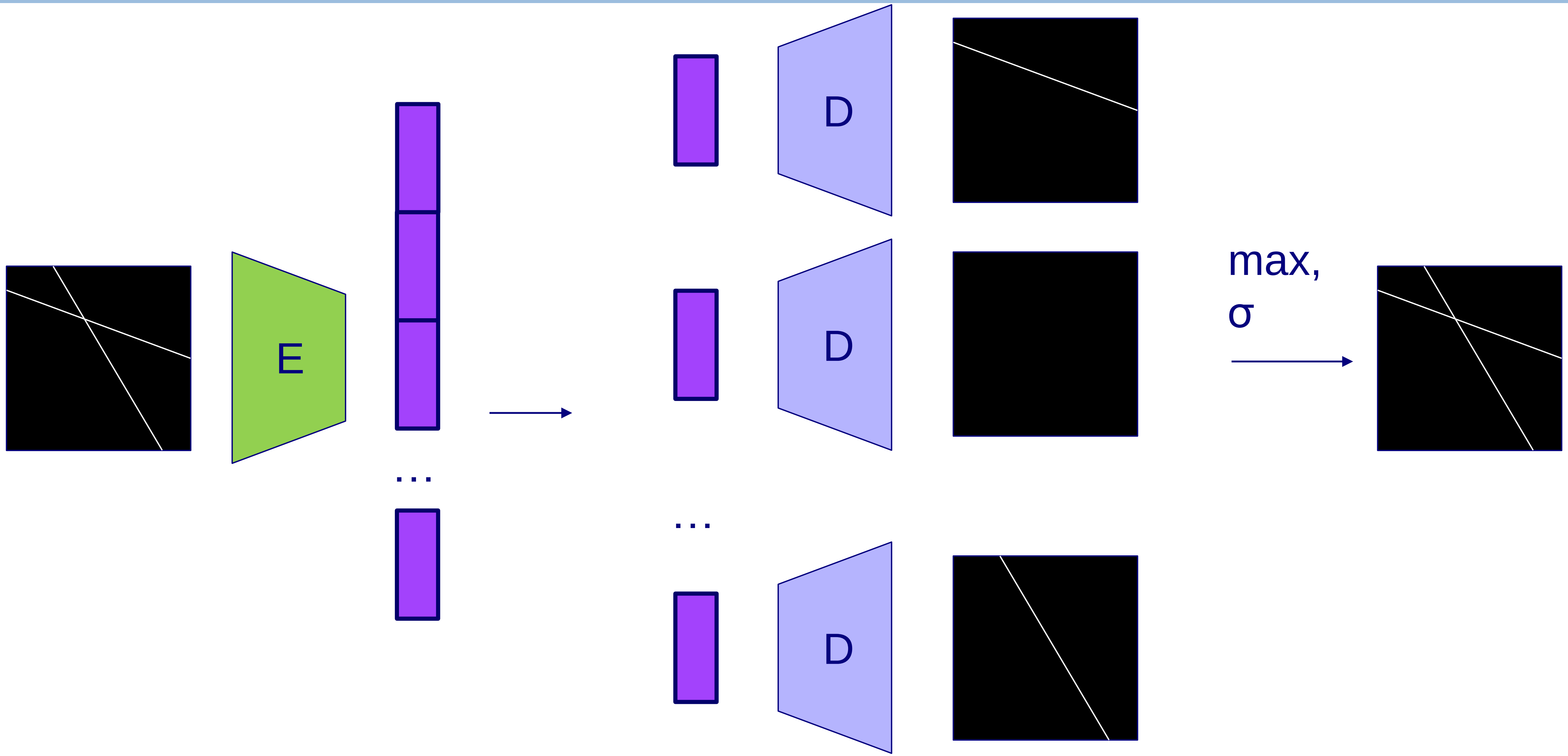


Transposed CoordConv: tiling, concatenation with x,y coordinate map, convolution



R. Liu et al. An intriguing failing of convolutional neural networks and the CoordConv solution. NIPS 2018

Model implementation: Autoencoder



Autoencoder performance

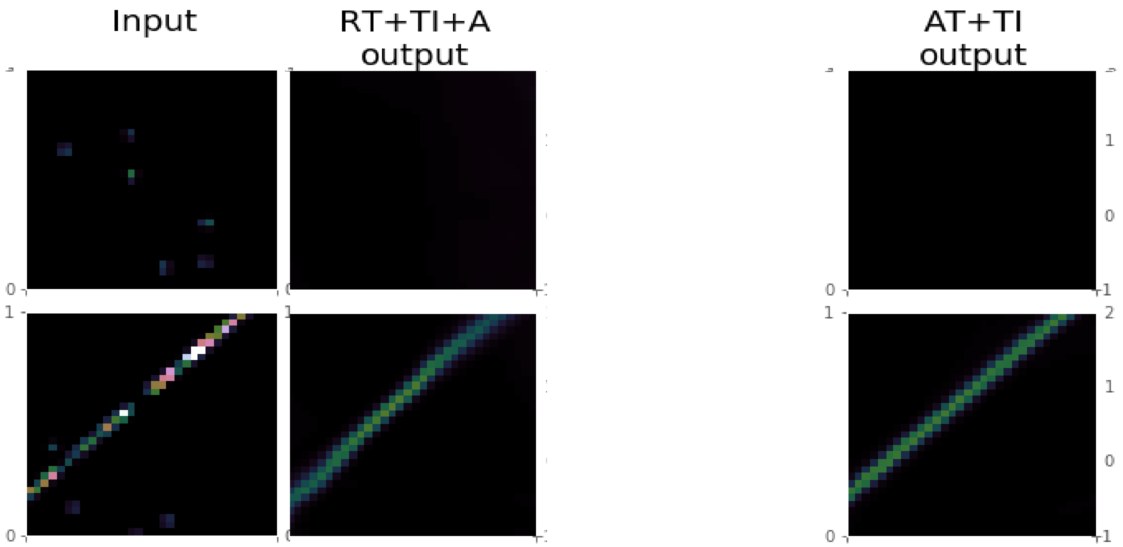
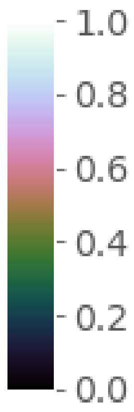
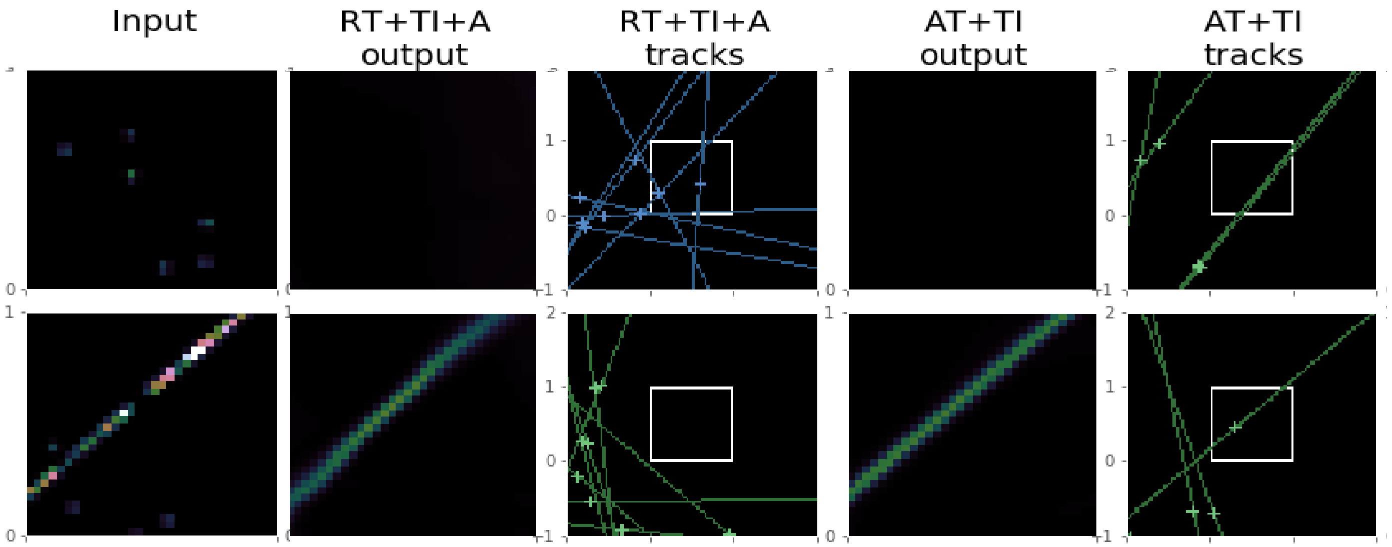


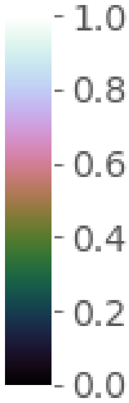
Image output is ok



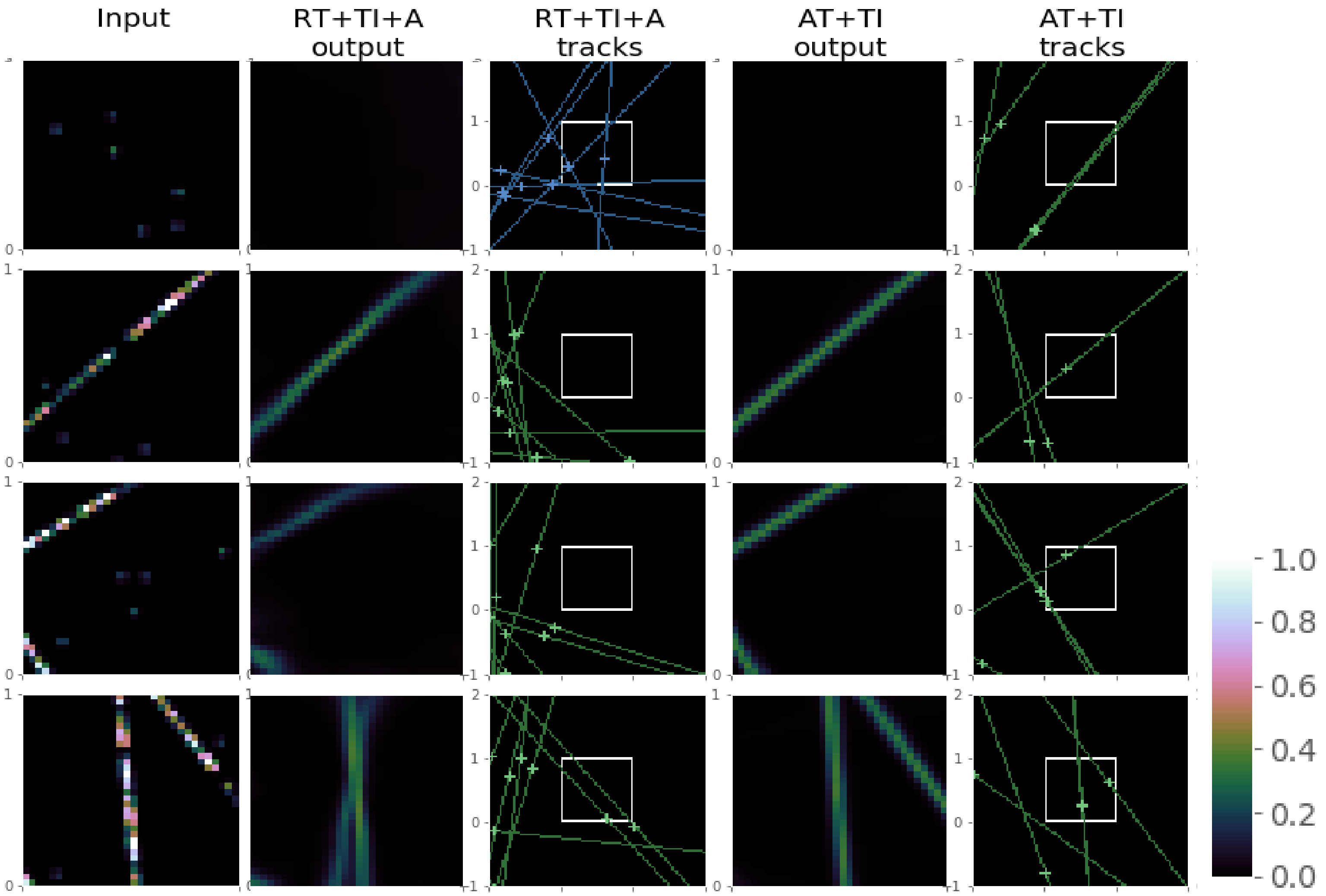
Autoencoder performance



AT+TI encodes line parameters properly!

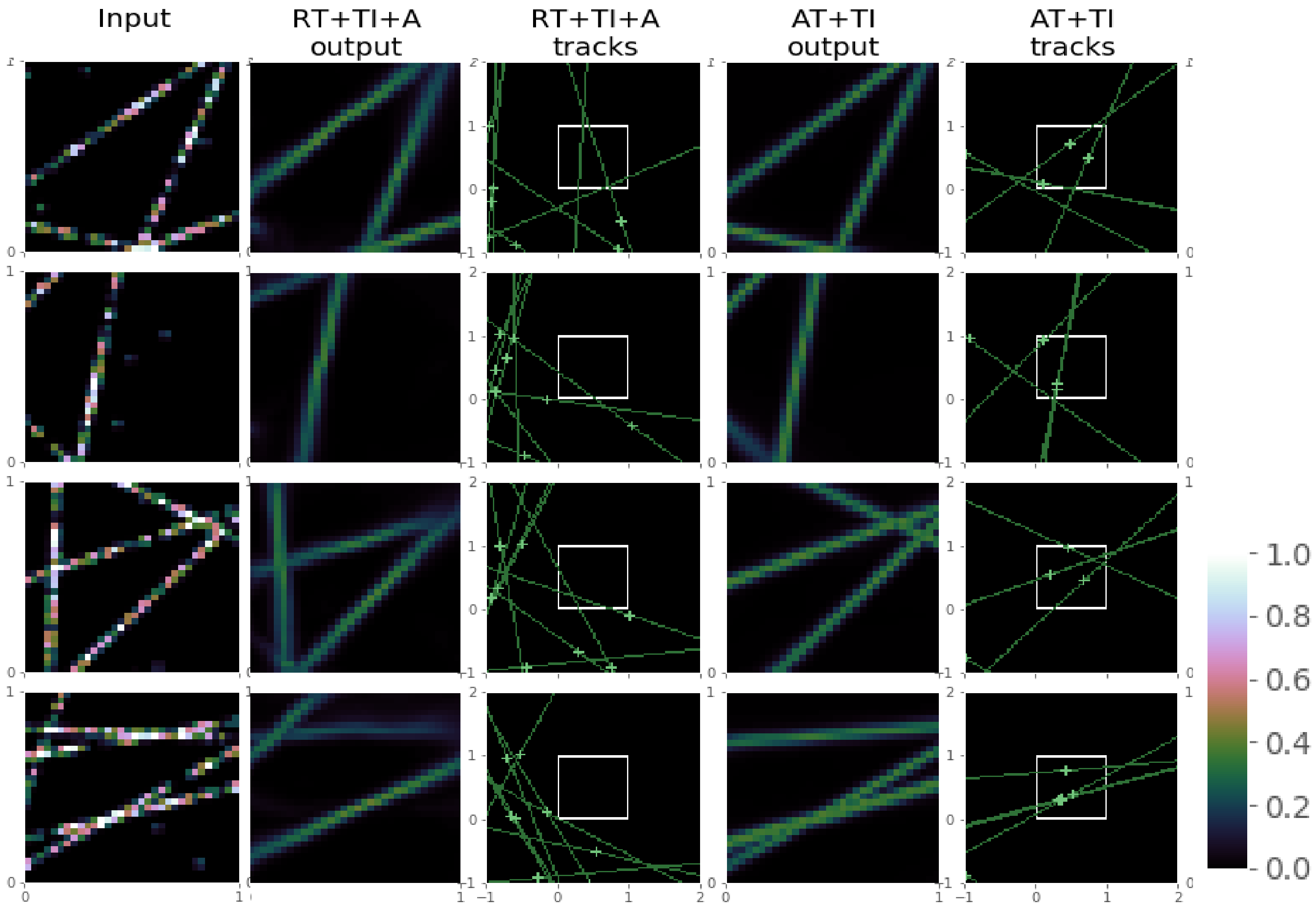


Autoencoder performance

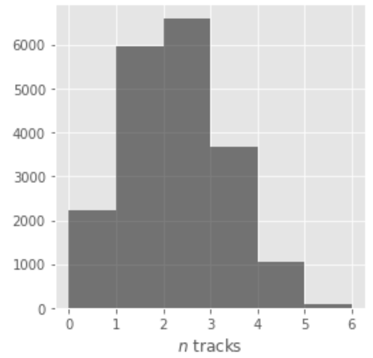


AT+TI encodes line parameters properly!

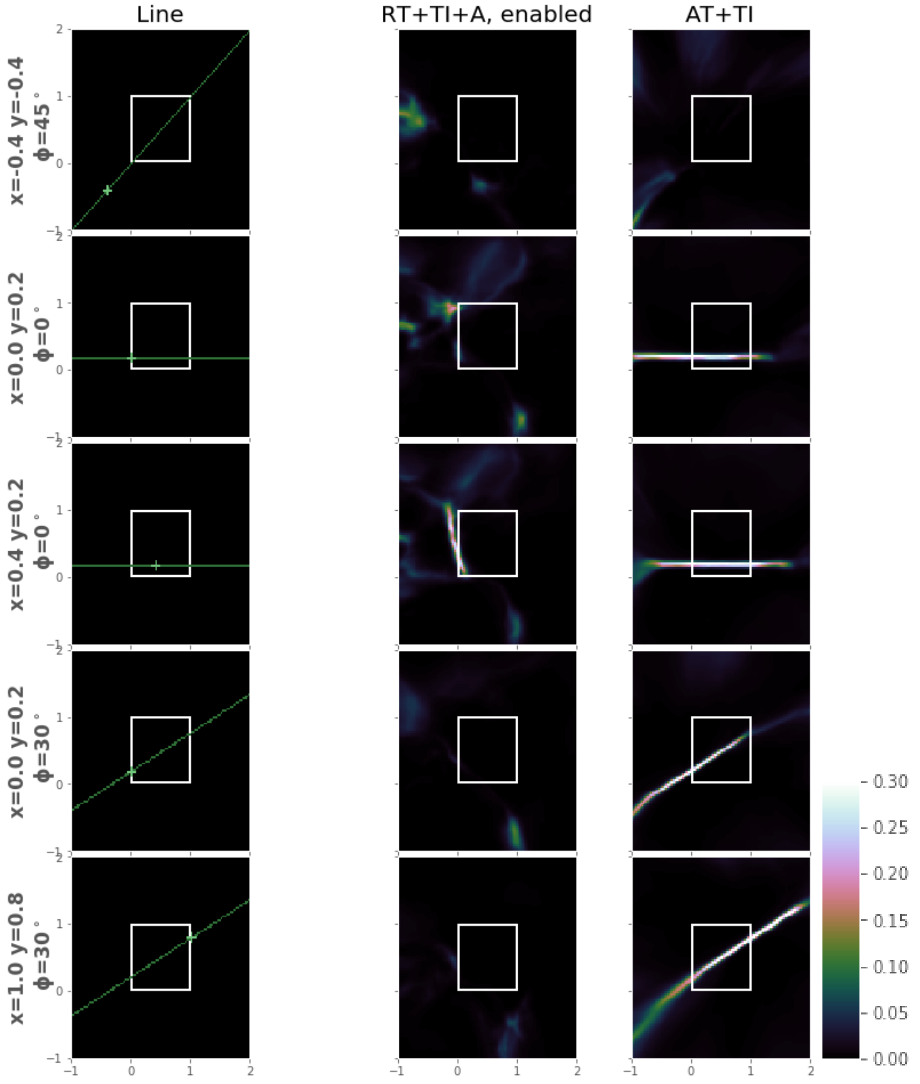
Autoencoder performance



Performance degrades with number of tracks



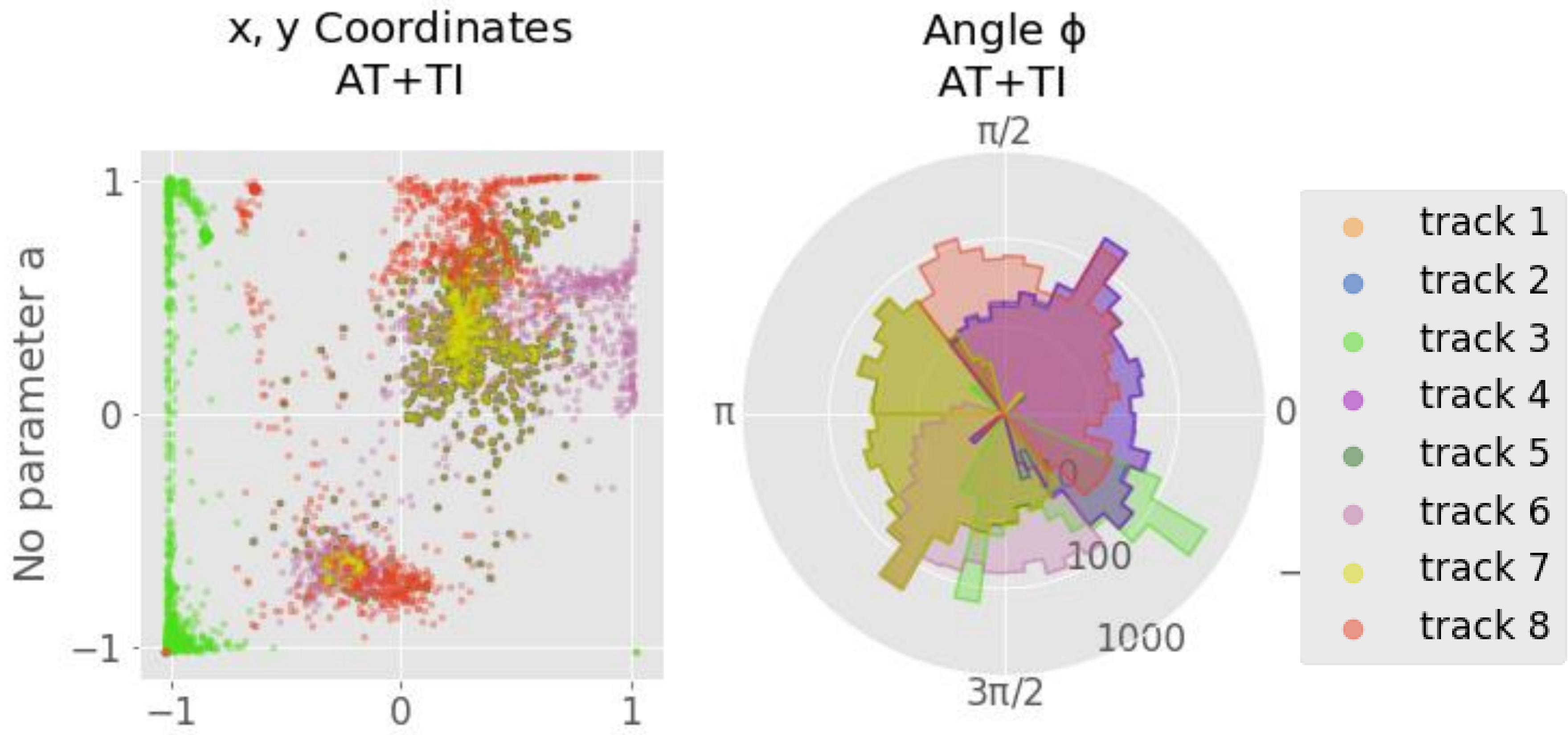
Learned disentanglement of the variational factors



Decoder output for given track parameters.

RT+TI did not learn desired reresentation

Latent parameter distribution



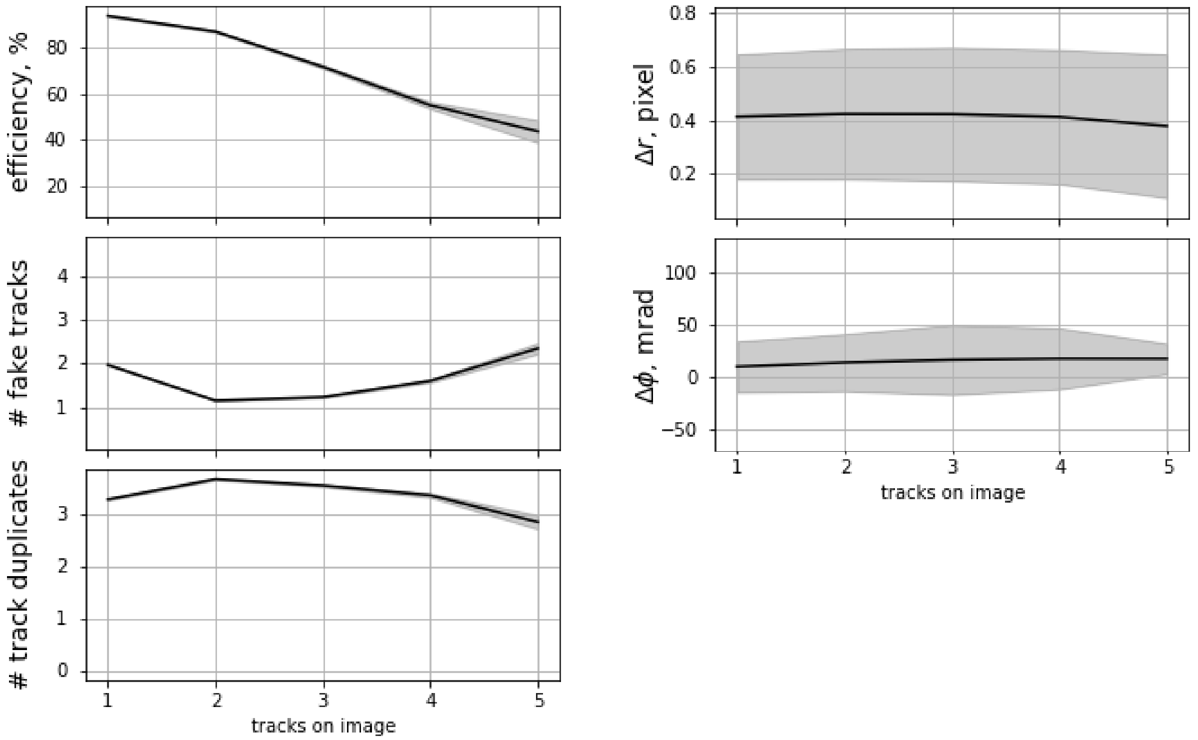
Track parameters’ measurement resolution

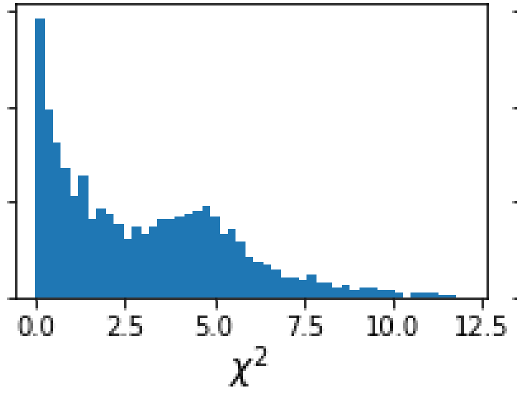
Processed 19600 test images

Assigned tracks as true, duplicates or fake by

$$\chi^2 = \left(\frac{\Delta r}{\sigma_r}\right)^2 + \left(\frac{\Delta \phi}{\sigma_\phi}\right)^2,$$

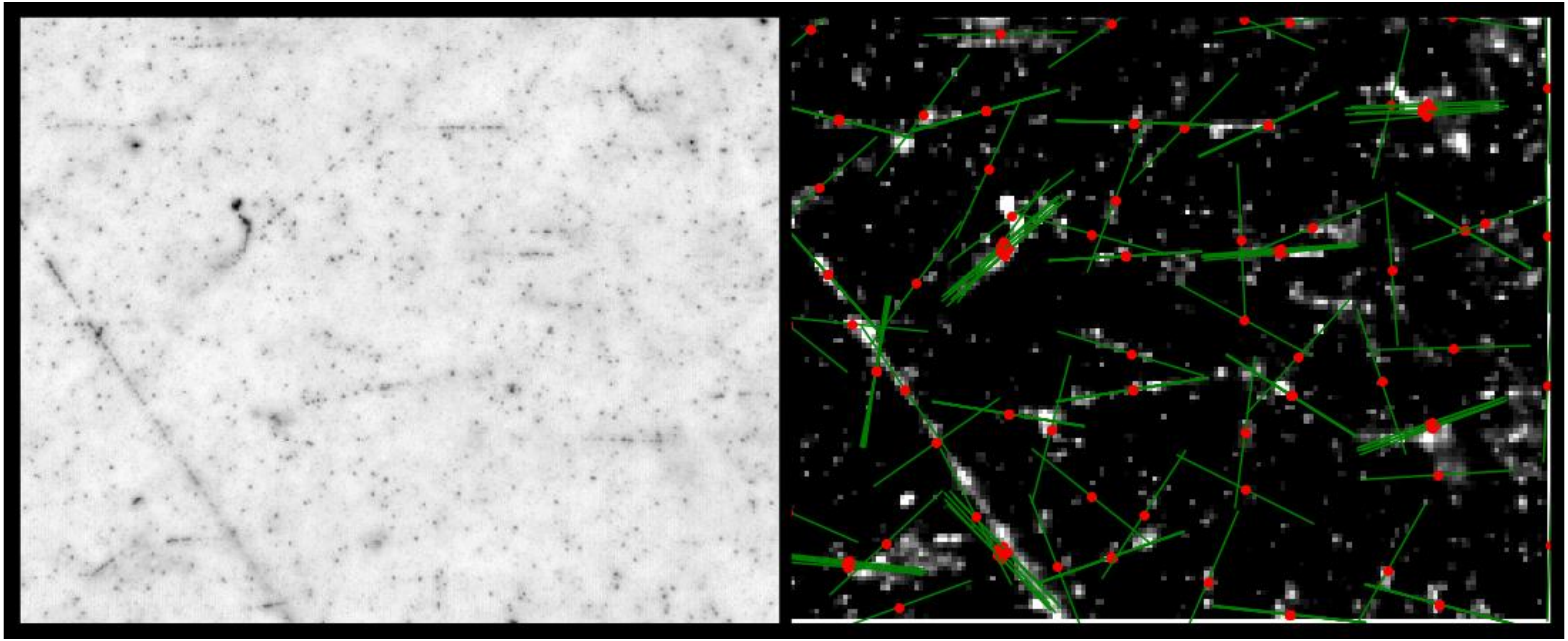
σ_r, σ_ϕ – max theoretical resolution





Gray bands: 1 sd of distribution for resolution parameters and of mean for efficiency, fake tracks and duplicates.

Emulsion data, AT+TI



Conclusion and outlook

- Imposing equivariance constraint on the autoencoder under geometrical transformations in the image and latent representation domains enables the model to “discover” the existence of multiple lines in the presented images in a fully unsupervised manner.

- Employing only the translation invariance (even together with rotation transformation) leads to reference ambiguity.

- Background and the grain distribution along the lines are yet to be integrated into the model.

M.Vladymyrov, A. Ariga,
Novel tracking approach based on fully-unsupervised disentanglement of the geometrical factors of variation

arXiv:1909.04427

Thank you!

PD Dr. A. Ariga

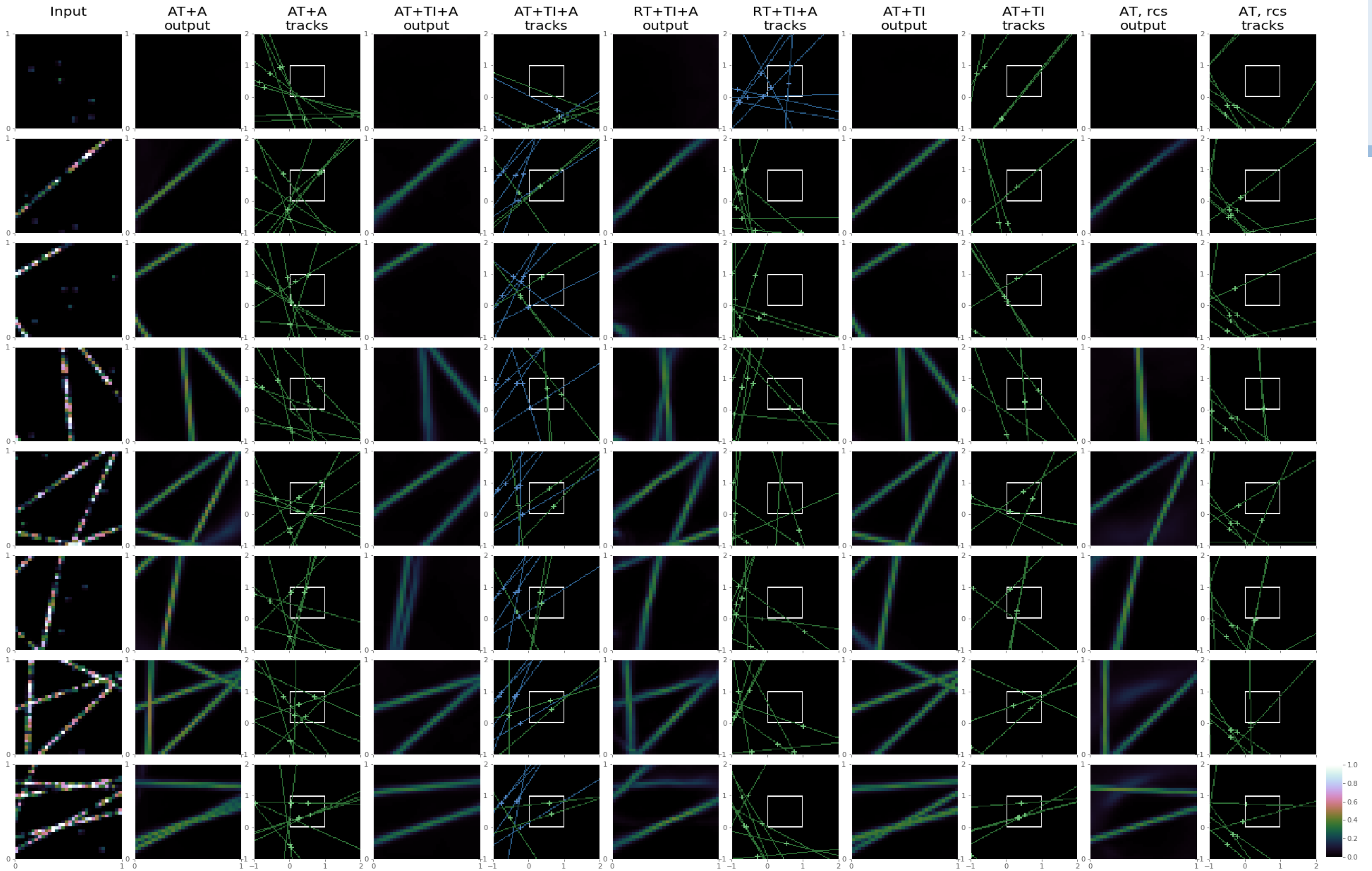
UNIVERSITÄT BERN

FNSNF

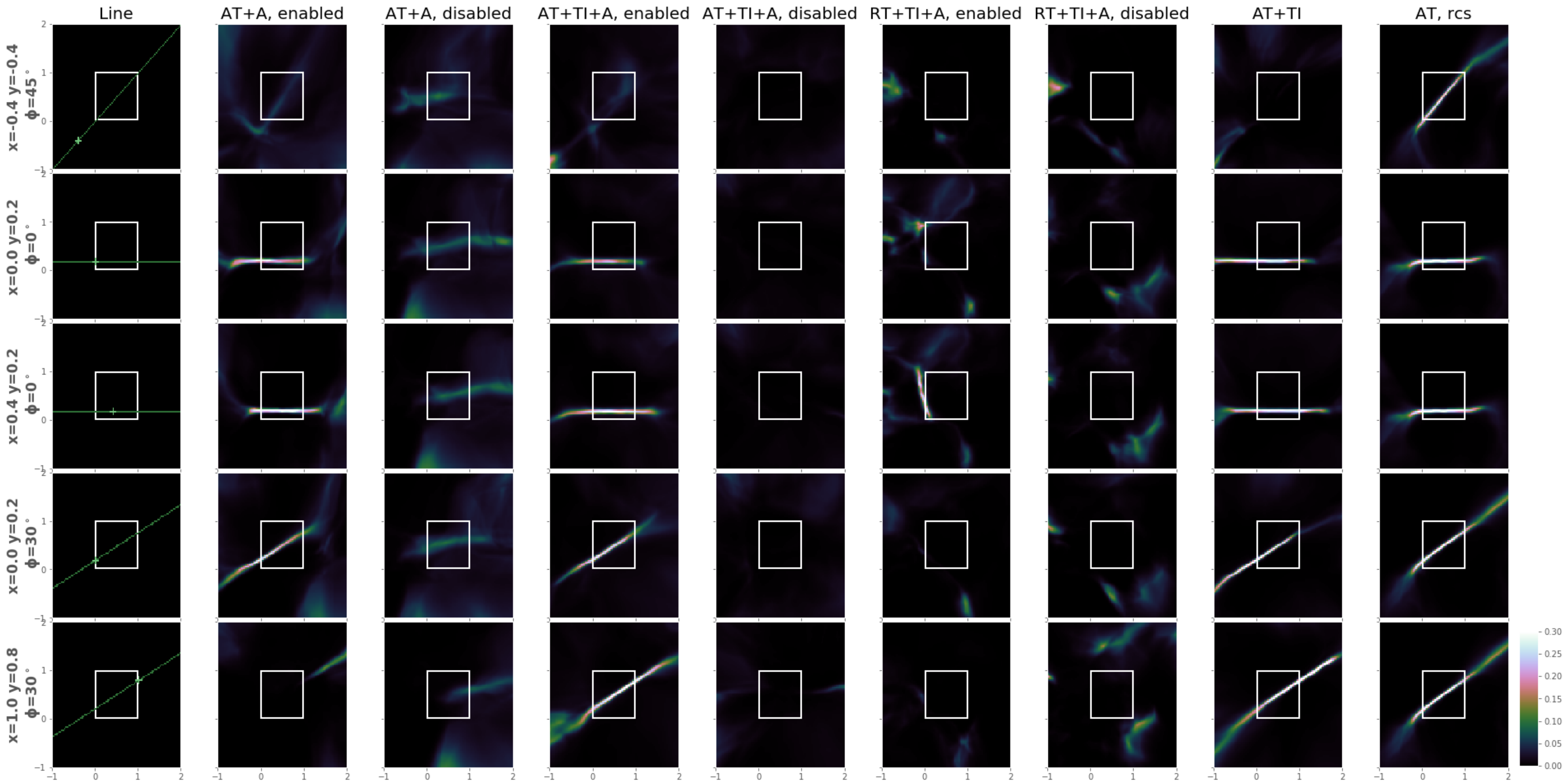
BACKUP SLIDES

Representation transformations

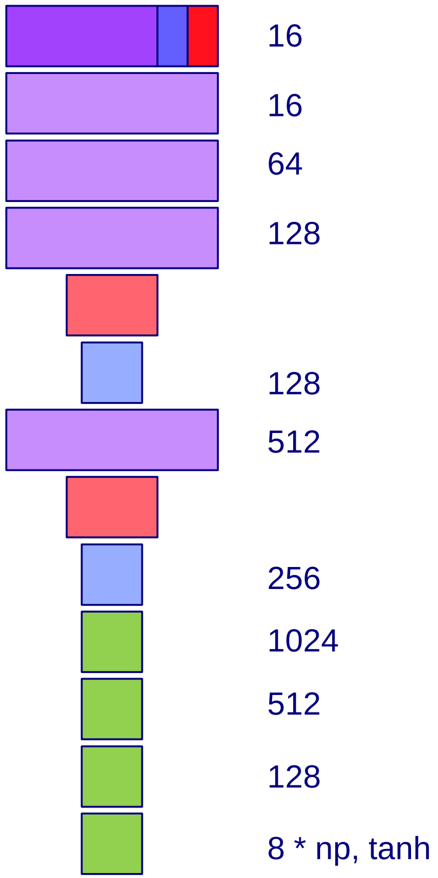
	Name	Description	n_p	Parametrization
1	AT+A	Affine Transformations with track activation parameter a	5	$(x_i, y_i, c_i, s_i, a_i)$
2	AT+TI+A	Affine Transformations + Translation Invariance with track activation parameter a	5	$(x_i, y_i, c_i, s_i, a_i)$
3	RT+TI+A	Translation Invariance + Rotation transformation only with track activation parameter a	5	$(x_i, y_i, c_i, s_i, a_i)$
4	AT+TI	Affine Transformations + Translation Invariance	4	(x_i, y_i, c_i, s_i)
5	AT, rcs	Affine Transformations in the (r, c, s) parametrization	3	(r_i, c_i, s_i)



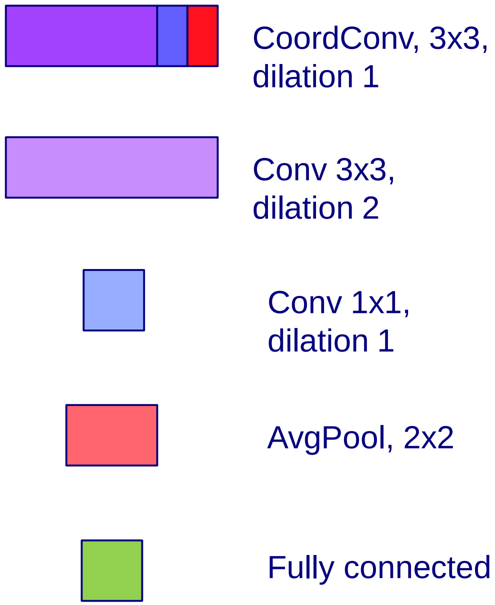
Input
AT+A output
AT+A tracks
AT+TI+A output
AT+TI+A tracks
RT+TI+A output
RT+TI+A tracks
AT+TI output
AT+TI tracks
AT, rcs output
AT, rcs tracks


Line
AT+A, enabled
AT+A, disabled
AT+TI+A, enabled
AT+TI+A, disabled
RT+TI+A, enabled
RT+TI+A, disabled
AT+TI
AT, rcs
x=-0.4 y=-0.4
ϕ=45°
x=0.0 y=0.2
ϕ=0°
x=0.4 y=0.2
ϕ=0°
x=0.0 y=0.2
ϕ=30°
x=1.0 y=0.8
ϕ=30°
0.30
0.25
0.20
0.15
0.10
0.05
0.00

Encoder

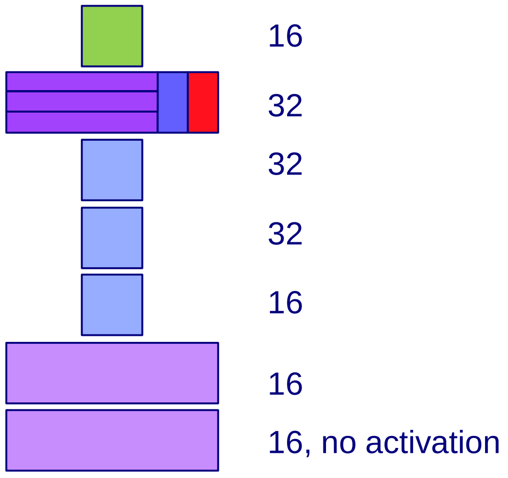


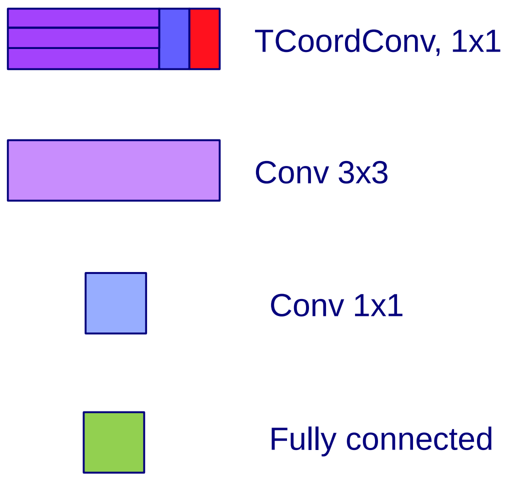
16
16
64
128
128
512
256
1024
512
128
8 * np, tanh



CoordConv, 3x3, dilation 1
Conv 3x3, dilation 2
Conv 1x1, dilation 1
AvgPool, 2x2
Fully connected

Decoder


16
32
32
32
16
16
16, no activation


TCoordConv, 1x1
Conv 3x3
Conv 1x1
Fully connected